

On the Quasi-Stationary Distribution of the Shiryaev–Roberts Diffusion

Aleksey S. Polunchenko

Department of Mathematical Sciences, State University of New York at Binghamton,
Binghamton, New York, USA

Abstract: We consider the diffusion $(R_t^r)_{t \geq 0}$ generated by the equation $dR_t^r = dt + \mu R_t^r dB_t$ with $R_0^r \triangleq r \geq 0$ fixed, and where $\mu \neq 0$ is given, and $(B_t)_{t \geq 0}$ is standard Brownian motion. We assume that $(R_t^r)_{t \geq 0}$ is stopped at $S_A^r \triangleq \inf\{t \geq 0: R_t^r = A\}$ with $A > 0$ preset, and obtain a closed-form formula for the quasi-stationary distribution of $(R_t^r)_{t \geq 0}$, i.e., the limit $Q_A(x) \triangleq \lim_{t \rightarrow +\infty} \mathbb{P}(R_t^r \leq x | S_A^r > t)$, $x \in [0, A]$. Further, we also prove $Q_A(x)$ to be unimodal for any $A > 0$, and obtain its entire moment series. More importantly, the pair (S_A^r, R_t^r) with $r \geq 0$ and $A > 0$ is the well-known Generalized Shiryaev–Roberts change-point detection procedure, and its characteristics for $r \sim Q_A(x)$ are of particular interest, especially when $A > 0$ is large. In view of this circumstance we offer an order-three large- A asymptotic approximation of $Q_A(x)$ valid for all $x \in [0, A]$. The approximation is rather accurate even if A is lower than what would be considered “large” in practice.

Keywords: Shiryaev–Roberts procedure; Quasi-stationary distribution; Quickest change-point detection; Whittaker functions.

Subject Classifications: 62L10; 60G40; 60J60.

1. INTRODUCTION

The general theme of this work is quickest change-point detection. The subject is concerned with the design and analysis of dependable “watchdog”-type statistical procedures for early detection of unanticipated changes that may (or may not) occur online in the characteristics of a “live”-monitored process. See, e.g., Shiryaev (1978), Basseville and Nikiforov (1993), Poor and Hadjiliadis (2009), Veeravalli and Banerjee (2013), and (Tartakovsky et al., 2014, Part II). A change-point detection procedure is a stopping time, T , that is adapted to the filtration, $(\mathcal{F}_t)_{t \geq 0}$, generated by the observed process, $(X_t)_{t \geq 0}$; the interpretation of T is that it is a rule to stop and declare that the statistical profile

Address correspondence to A. S. Polunchenko, Department of Mathematical Sciences, State University of New York (SUNY) at Binghamton, 4400 Vestal Parkway East, Binghamton, NY 13902–6000, USA; Tel: +1 (607) 777–6906; Fax: +1 (607) 777–2450; E-mail: aleksey@binghamton.edu.

of the observed process may have (been) changed. A “good” (i.e., optimal or nearly optimal) detection procedure is one that minimizes (or nearly minimizes) the desired detection delay penalty-function, subject to a constraint on the “false alarm” risk. For an overview of the major optimality criteria see, e.g., Tartakovsky and Moustakides (2010), Polunchenko and Tartakovsky (2012), Veeravalli and Banerjee (2013), and (Tartakovsky et al., 2014, Part II).

This work is motivated by the classical minimax change-point detection problem where the observed process, $(X_t)_{t \geq 0}$, is standard Brownian motion that at an unknown (nonrandom) time moment $\nu \geq 0$ —referred to as the change-point—may (or may not) experience an abrupt and permanent change in the drift, from a value of zero initially, i.e., $\mathbb{E}[X_t] = 0$ for $t \in [0, \nu]$, to a known value $\mu \neq 0$ following the change-point, i.e., $\mathbb{E}[X_t] = \mu t$ for $t > \nu$. The goal is to find out—as quickly as is possible within an *a priori* set level of the “false positive” risk—whether the drift of the process is no longer zero. See, e.g., Pollak and Siegmund (1985), Shiryaev (1996, 2002), Moustakides (2004), Shiryaev (2006), Feinberg and Shiryaev (2006), Burnaev et al. (2009), (Shiryaev, 2011, Chapter 5), Polunchenko and Sokolov (2016) and Polunchenko (2016). More formally, under the Brownian motion change-point scenario, the observed process, $(X_t)_{t \geq 0}$, is governed by the stochastic differential equation (SDE):

$$dX_t = \mu \mathbb{1}_{\{t > \nu\}} dt + dB_t, \quad t \geq 0, \quad \text{with } X_0 = 0, \quad (1.1)$$

where $(B_t)_{t \geq 0}$ is standard Brownian motion (i.e., $\mathbb{E}[dB_t] = 0$, $\mathbb{E}[(dB_t)^2] = dt$, and $B_0 = 0$), $\mu \neq 0$ is the *known* post-change drift value, and $\nu \in [0, \infty]$ is the unknown (nonrandom) change-point; here and onward, the notation $\nu = 0$ ($\nu = \infty$) is to be understood as the case when the drift is in effect *ab initio* (or never, respectively). Let \mathbb{P}_∞ (\mathbb{P}_0) denote the probability measure (distribution law) generated by the observed process, $(X_t)_{t \geq 0}$, under the assumption that $\nu = \infty$ ($\nu = 0$); note that \mathbb{P}_∞ is the Wiener measure.

A sensible way to perform change-point detection under model (1.1) is to use the Generalized Shiryaev–Roberts (GSR) procedure proposed by Moustakides et al. (2011) as a headstarted (hence, more general) version of the classical quasi-Bayesian Shiryaev–Roberts (SR) procedure that emerged from the independent work of Shiryaev (1961; 1963) and that of Roberts (1966). Specifically, tailored to the Brownian motion scenario (1.1), the GSR procedure is given by the stopping time:

$$\mathcal{S}_A^r \triangleq \inf\{t \geq 0: R_t^r = A\} \text{ such that } \inf\{\emptyset\} = \infty, \quad (1.2)$$

where $A > 0$ is the detection threshold (set in advance in such a way that the “false positive” risk is within the desired margin of tolerance), and the GSR statistic $(R_t^r)_{t \geq 0}$ is the diffusion process that satisfies the stochastic differential equation

$$dR_t^r = dt + \mu R_t^r dX_t \text{ with } R_0^r \triangleq r \geq 0, \quad (1.3)$$

with $(X_t)_{t \geq 0}$ as in (1.1); the initial value r is known as the headstart.

The choice to go with the GSR procedure may be justified by the result previously obtained (for the discrete-time analogue of the problem) by Tartakovsky et al. (2012) where the GSR procedure with a carefully designed headstart was shown to be nearly minimax in the sense of Pollak (1985); an attempt to generalize this result to the Brownian

motion scenario (1.1) was made, e.g., by Burnaev (2009). The GSR procedure's near Pollak-minimaxity is a strong optimality property known in the literature as *order-three* minimaxity because the respective delay cost is minimized up to an additive term that goes to zero together with the “false alarm” risk level. Also, it was demonstrated explicitly by Tartakovsky and Polunchenko (2010) and by Polunchenko and Tartakovsky (2010) that in two specific (discrete-time) scenarios the GSR procedure (again with a “finetuned” headstart) is actually *exactly* minimax in Pollak's (1985) sense. More importantly, while a *general* solution to Pollak's (1985) minimax change-point detection problem is still unknown, there is a universal “recipe” (also proposed by Pollak 1985) to achieve strong order-three Pollak-minimaxity, and the main ingredient of the “recipe” is the GSR procedure. Specifically, the idea is to start the GSR statistic off a random number sampled from the statistic's so-called quasi-stationary distribution (formally defined below). Pollak (1985) proved that, in the discrete-time setup, such a randomized “tweak” of the GSR procedure is order-three Pollak-minimax; see also Tartakovsky et al. (2012). The same may well hold true for the Brownian motion scenario (1.1) too, although, to the best of our knowledge, this question has not yet been investigated in the literature, except in the work of Burnaev et al. (2009) where the GSR procedure was shown to be almost Pollak-minimax, but only up to the second order (the delay cost is minimized up to an additive term that goes to a constant as the “false alarm” risk level goes to zero). Even though Pollak's (1985) idea to randomize the GSR procedure as described above *does not* necessarily lead to *strict* Pollak-minimaxity (see Polunchenko and Tartakovsky 2010; Tartakovsky and Polunchenko 2010 for counterexamples), it does allow to achieve order-three minimaxity in the discrete-time setup, and whether or not this is also the case for the continuous-time scenario (1.1) is a problem that is certainly worthy of consideration. This work is an attempt to make the first step in this direction. Incidentally, the quasi-stationary distribution is also essential for the evaluation of the GSR procedure's minimax performance in the stationary regime, i.e., when $\nu \rightarrow +\infty$; see Pollak and Siegmund (1985).

More concretely, the overall aim of this paper is to obtain an exact closed-form formula for the GSR statistic's (1.3) quasi-stationary distribution and an accurate asymptotic approximation thereof for when the GSR procedure's detection threshold $A > 0$ is large. Formally, the sought quasi-stationary distribution is defined as

$$Q_A(x) \triangleq \lim_{t \rightarrow +\infty} \mathbb{P}_\infty(R_t^r \leq x | \mathcal{S}_A^r > t) \text{ with } q_A(x) \triangleq \frac{d}{dx} Q_A(x), \quad (1.4)$$

which is independent of the headstart $R_0^r \triangleq r \geq 0$, provided $r \in [0, A]$, i.e., within the support of the density $q_A(x)$. The existence of this distribution for the Brownian motion scenario (1.1) follows from the results previously obtained in the fundamental work of Mandl (1961); see also, e.g., Cattiaux et al. (2009). While we do get $Q_A(x)$ and $q_A(x)$ expressed analytically and explicitly (see Section 3), the obtained formulae involve special functions, and, as a result, a *precise* performance analysis of Pollak's (1985) randomized GSR procedure is problematic: the calculus involved is prohibitively difficult. The usual way around this is to assume the GSR procedure's detection threshold $A > 0$ is large and study the randomized GSR procedure asymptotically as $A \rightarrow +\infty$. To that end, it is known (see Pollak and Siegmund 1985, 1986) that $Q_A(x)$ does have a limit as $A \rightarrow +\infty$, and the limit is the GSR statistic's so-called stationary distribution defined as

$$H(x) \triangleq \lim_{t \rightarrow +\infty} \mathbb{P}_\infty(R_t^r \leq x) \text{ with } h(x) \triangleq \frac{d}{dx} H(x), \text{ where } x \in [0, +\infty), \quad (1.5)$$

which is also independent of the headstart $R_0^r \triangleq r \geq 0$; the convergence of $Q_A(x)$ to $H(x)$ as $A \rightarrow +\infty$ is pointwise, at all continuity points of $H(x)$. Although the stationary distribution $H(x)$ has already been found and studied in the literature (see, e.g., Shiryayev 1961, 1963, Pollak and Siegmund 1985, Feinberg and Shiryayev 2006, or Burnaev et al. 2009), a “little oh”-level investigation of the randomized GSR procedure’s characteristics requires a more “fine” large- A asymptotic approximation of the quasi-stationary distribution itself. To that end, we offer a large- A *order-three* expansion of the density $q_A(x)$ valid for all $x \in [0, A]$. The expansion is derived directly from the exact formula for $q_A(x)$, with the aid of the Mellin integral transform, which, if need be, can also be used to expand $q_A(x)$ even further, beyond the third-order term. As an auxiliary result, we prove that the quasi-stationary distribution is unimodal for any $A > 0$; it is of note that its limit as $A \rightarrow +\infty$, i.e., the stationary distribution (1.5), is known to be unimodal as well. We also obtain the quasi-stationary distribution’s entire moment series.

The remainder of the paper is organized as follows. Section 2 fixes nomenclature and notation, and provides the necessary preliminary background. The main contribution—i.e., the quasi-stationary distribution and its properties—is the subject of Section 3. Conclusions and outlook follow in Section 4. Appendix offers proofs of certain lemmas.

2. PRELIMINARIES

For notational brevity, we shall henceforth omit the subscript “ A ” in “ $Q_A(x)$ ” as well as in “ $q_A(x)$ ”, unless the dependence on A is noteworthy. Also, for technical convenience and without loss of generality, we shall primarily deal with $q(x)$ rather than with $Q(x)$.

It has already been established in the literature (see, e.g., Mandl 1961, Cattiaux et al. 2009, or Burnaev et al. 2009, Equation (35), p. 528) that $q(x)$, formally defined in (1.4), is the solution of a certain boundary-value problem composed of a second-order ordinary differential equation (ODE) considered on the interval $[0, A]$, a pair of boundary conditions, and a normalization constraint. Specifically, the ODE—which we shall henceforth refer to as the *master equation*—is of the form

$$\frac{\mu^2}{2} \frac{d^2}{dx^2} [x^2 q(x)] - \frac{d}{dx} [q(x)] = \lambda q(x), \quad x \in [0, A], \quad (2.1)$$

where $\lambda^{(0)}$ is the *dominant* eigenvalue of the differential operator

$$\mathcal{D} \triangleq \frac{\mu^2}{2} \frac{\partial^2}{\partial x^2} x^2 - \frac{\partial}{\partial x}, \quad (2.2)$$

i.e., the infinitesimal generator of the GSR diffusion $(R_t^r)_{t \geq 0}$ under the \mathbb{P}_∞ probability measure; observe from (1.1) and (1.3) that the \mathbb{P}_∞ -differential of the GSR diffusion $(R_t^r)_{t \geq 0}$ is $dR_t^r = dt + \mu R_t^r dB_t$ where $R_0^r \triangleq r \in [0, A]$ and $A > 0$. It goes without saying that λ (or any other eigenvalue of the operator \mathcal{D} for that matter) is dependent on A , and, wherever necessary, we shall emphasize this dependence via the notation λ_A .

Next, the behavior of $q(x)$ near the left end-point of the interval $[0, A]$ —which is the GSR diffusion’s domain—must be such that

$$\lim_{x \rightarrow 0+} \left\{ \frac{\mu^2}{2} \frac{\partial}{\partial x} [x^2 q(x)] - q(x) \right\} = 0, \quad (2.3)$$

which is to say that $x = 0$ is, in Feller's (1952) classification, an *entrance* boundary for the process $(R_t^r)_{t \geq 0}$; in “differential equations speak”, this is a Neumann-type boundary condition. The boundary condition at the other end of the interval $[0, A]$ is of the form

$$q(A) = 0, \quad (2.4)$$

i.e., the density $q(x)$ must vanish at $x = A$, for, by definition (1.2) of the GSR stopping time, the GSR diffusion is “killed” or “absorbed” once it hits the detection threshold $A > 0$; in “differential equations speak”, this is a Dirichlet-type boundary condition.

Lastly, the density $q(x)$ must also satisfy the obvious normalization constraint

$$\int_0^A q(x) dx = 1. \quad (2.5)$$

Subject to boundary conditions (2.3) and (2.4), equation (2.1) is a Sturm–Liouville problem. It is straightforward to see that by virtue of the multiplying factor

$$\mathbf{m}(x) \triangleq \frac{2}{\mu^2 x^2} e^{-\frac{2}{\mu^2 x}} \quad (2.6)$$

the equation can be brought to the canonical Sturm–Liouville form

$$\frac{\mu^2}{2} \frac{d}{dx} \left[x^2 \mathbf{m}(x) \frac{d}{dx} \varphi(x) \right] = \lambda \mathbf{m}(x) \varphi(x), \quad (2.7)$$

where the unknown function $\varphi(x)$ is such that $q(x) \propto \mathbf{m}(x) \varphi(x)$, i.e., $q(x)$ is a multiple of $\mathbf{m}(x) \varphi(x)$. Hence, our problem effectively is to consider the Sturm–Liouville operator

$$\mathcal{G} \triangleq \frac{\mu^2}{2 \mathbf{m}(x)} \frac{d}{dx} x^2 \mathbf{m}(x) \frac{d}{dx} \quad (2.8)$$

with $\mathbf{m}(x)$ given by (2.6), restrict it to the interval $x \in [0, A]$, and recover its dominant eigenvalue $\lambda \equiv \lambda_A$ and the respective eigenfunction $\varphi(x) \equiv \varphi(x, \lambda)$ for which

$$\lim_{x \rightarrow 0+} \left[x^2 \mathbf{m}(x) \frac{d}{dx} \varphi(x, \lambda) \right] = 0 \quad \text{and} \quad \varphi(A, \lambda) = 0. \quad (2.9)$$

The general theory of second-order differential operators or Sturm–Liouville operators (such as our operators \mathcal{D} and \mathcal{G} defined above) is well-developed, and, in particular, the spectral properties of such operators are well-understood. The classical fundamental references on the subject are Titchmarsh (1962), Levitan (1950), Coddington and Levinson (1955), Dunford and Schwartz (1963), and Levitan and Sargsjan (1975); for applications of the theory to diffusion processes, see, e.g., (Itô and McKean, 1974, Section 4.11), and Linetsky (2007) who provides a great overview of the state-of-the-art in the field. We now recall a few results from the general Sturm–Liouville theory that directly apply to our specific Sturm–Liouville problem. These results will be utilized in the sequel, and, conveniently enough, all of them (and much more) can be found in the work of Fulton et al. (1999); see also, e.g., Kent (1980).

We start with (Fulton et al., 1999, Theorem 18, p. 22) which asserts that our Sturm–Liouville problem given by equation (2.7) and two boundary conditions (2.9) does have a solution $\varphi(x, \lambda)$ defined up to an arbitrary multiplicative

factor independent of x . Consequently, by invoking the normalization constraint (2.5), one can conclude that the quasi-stationary pdf $q(x)$ does exist and is unique.

Next, we turn to (Fulton et al., 1999, Section 7) which introduces ten mutually exclusive categories to classify Sturm–Liouville operators depending on their spectral properties. Our Sturm–Liouville problem belongs to Spectral Category 1 (see Fulton et al. 1999, p. 22). This means, among many things, that the spectrum of the operator \mathcal{D} given by (2.2) is purely discrete, and is determined entirely by the Dirichlet boundary condition (2.4). More concretely, the spectrum is the set of solutions of the equation $\varphi(A, \lambda) = 0$ where λ is the unknown and $A > 0$ is fixed. This equation has countably many distinct zeros, say $\{\lambda^{(i)}\}_{i \geq 0}$, and each one of them is simple, and their series for any fixed $0 < A < +\infty$ is such that $0 < -\lambda^{(0)} < -\lambda^{(1)} < -\lambda^{(2)} < \dots$ with $-\lim_{i \rightarrow +\infty} \lambda^{(i)} = +\infty$. Furthermore, for the dominant eigenvalue $\lambda_A (\equiv \lambda^{(0)})$ the equation $\varphi(x, \lambda_A) = 0$ with x being the unknown has no zeros inside the interval $x \in (0, A)$. All this translates to the following two results.

Lemma 2.1. *For any fixed $A > 0$ the quasi-stationary pdf $q_A(x)$ has no zeros inside the interval $(0, A)$, i.e., $q_A(x) > 0$ for all $x \in (0, A)$.*

Lemma 2.2. *For any fixed $A > 0$ the dominant eigenvalue $\lambda \equiv \lambda_A$ of the operator \mathcal{D} given by (2.2) is nonpositive, i.e., $\lambda_A \leq 0$.*

Proof. The claim is actually true for any eigenvalue of the operator \mathcal{D} . For an explicit proof adapted to the context of diffusion processes, see, e.g., (Wong, 1964, Section 1, pp. 265–266) or (Polunchenko, 2016, Section 3, pp. 125–126). \square

It is of note that $\lambda = 0$ and $\varphi(x, 0) \equiv 1$ is always a (nontrivial) solution of (2.7), and $q(x)$ in this case is a multiple of $m(x)$ given by (2.6). However, while $q(x) \propto m(x)$ does (trivially) satisfy the entrance boundary condition (2.3) for any A , the absorbing boundary condition (2.4) is satisfied only in the limit as $A \rightarrow +\infty$. Moreover, since $m(x) \geq 0$ for all $x \in \mathbb{R}$ and because it integrates to unity over $x \in [0, +\infty)$, the normalization constraint (2.5) is automatically fulfilled for $q(x) = m(x)$ in the limit as $A \rightarrow +\infty$. We therefore arrive at the result first obtained by Shiryayev (1961; 1963) that $q(x) = m(x)$ with $x \in [0, +\infty)$ is the stationary distribution (1.5) of the GSR statistic, i.e., $h(x) \triangleq \lim_{A \rightarrow +\infty} q_A(x) = m(x)$, $x \in [0, +\infty)$; see also, e.g., Pollak and Siegmund (1985), Feinberg and Shiryayev (2006), (Burnaev et al., 2009, Remark 2, p. 529), or Polunchenko and Sokolov (2016). This momentless distribution is a special case of the (extreme-value) Fréchet distribution.

We conclude this section by recalling that, according to the well-known basic result (see, e.g., Levitan 1950, Lemma 1.1.1 or Levitan and Sargsjan 1975, Lemma 1.1, pp. 2–3) from the general Sturm–Liouville theory, the eigenfunctions $\varphi(x, \lambda)$ indexed by λ of the operator \mathcal{D} form an orthonormal basis in the Hilbert space $L^2([0, A], m)$ of real-valued “ $m(x)$ ”-measurable, square-integrable (with respect to the “ $m(x)$ ”-measure) functions defined on the interval $[0, A]$ equipped with the “ $m(x)$ ”-weighted inner product:

$$\langle f, g \rangle_m \triangleq \int_0^A m(x) f(x) g(x) dx.$$

More specifically, the foregoing means that if $\varphi(x, \lambda^{(i)})$ and $\varphi(x, \lambda^{(j)})$ are any two eigenfunctions of the operator \mathcal{G} given by (2.8), then

$$\int_0^A \mathfrak{m}(x) \varphi(x, \lambda^{(i)}) \varphi(x, \lambda^{(j)}) dx = \mathbb{1}_{\{i=j\}},$$

where $\varphi(x, \lambda^{(i)})$ and $\varphi(x, \lambda^{(j)})$ are each assumed to be of unit “length”, i.e., $\|\varphi(\cdot, \lambda^{(i)})\| = 1 = \|\varphi(\cdot, \lambda^{(j)})\|$, with the “length” defined as

$$\|\varphi(\cdot, \lambda)\|^2 \triangleq \int_0^A \mathfrak{m}(x) \varphi^2(x, \lambda) dx. \quad (2.10)$$

We are now in a position to attack the master equation (2.1) directly. This is the object of the next section, which is the main section of this paper.

3. THE QUASI-STATIONARY DISTRIBUTION AND ITS CHARACTERISTICS

The plan now is to first fix the detection threshold $A > 0$ and solve the master equation (2.1) analytically to recover $q_A(x)$ and $Q_A(x)$, both in a closed form and for all $x \in [0, A]$, and then assume that A is large and consider the asymptotic (as $A \rightarrow +\infty$) behavior of $q_A(x)$ on its support $x \in [0, A]$. Much of the exact analytical solution part is a build-up to the closely related earlier work of the author (2016); see also Linetsky (2004) and Polunchenko and Sokolov (2016).

3.1. The Exact Formulae and Analysis

The idea is to apply the change of variables

$$x \mapsto u = u(x) = \frac{2}{\mu^2 x}, \text{ so that } u \mapsto x = x(u) = \frac{2}{\mu^2 u} \text{ and } \frac{dx}{x} = -\frac{du}{u}, \quad (3.1)$$

along with the substitution

$$\varphi(x) \mapsto \varphi(u) \triangleq \frac{v(u)}{\sqrt{\mathfrak{m}(u)}} = \sqrt{\frac{\mu^2}{2}} \frac{1}{u} e^{\frac{u}{2}} v(u) \propto \frac{1}{u} e^{\frac{u}{2}} v(u), \quad (3.2)$$

to bring our equation to the form

$$v_{uu}(u) + \left\{ -\frac{1}{4} + \frac{1}{u} + \frac{1/4 - \xi^2/4}{u^2} \right\} v(u) = 0, \quad (3.3)$$

where

$$\xi \equiv \xi(\lambda) \triangleq \sqrt{1 + \frac{8}{\mu^2} \lambda} \text{ so that } \lambda \equiv \lambda(\xi) = \frac{\mu^2}{8} (\xi^2 - 1), \quad (3.4)$$

and one may see that ξ is either purely real (ranging between 0 and 1) or purely complex, because $\lambda \leq 0$, as asserted by Lemma 2.2 above. We also remark that equation (3.3) is symmetric with respect to the sign of ξ .

The obtained equation (3.3) is a particular version of the classical Whittaker Whittaker (1904) equation

$$w_{zz}(z) + \left\{ -\frac{1}{4} + \frac{a}{z} + \frac{1/4 - b^2}{z^2} \right\} w(z) = 0, \quad (3.5)$$

where $w(z)$ is the unknown function of $z \in \mathbb{C}$, and $a, b \in \mathbb{C}$ are two given parameters; see, e.g., (Buchholz, 1969, Chapter I). A self-adjoint homogeneous second-order ODE, Whittaker's equation (3.5) provides a definition for the well-known two Whittaker functions: the latter may be defined as the equation's two fundamental solutions. More concretely, the two Whittaker functions are special functions that are conventionally denoted $W_{a,b}(z)$ and $M_{a,b}(z)$, where the indices a and b are the Whittaker's equation (3.5) parameters. The functions are typically considered in the cut plane $|\arg(z)| < \pi$ to ensure they are not multi-valued. For an extensive study of these functions and various properties thereof, see, e.g., Buchholz (1969) or Slater (1960).

By combining (3.5), (3.3), (3.2) and (3.1) one can see that the *general* form of $\varphi(x, \lambda)$ is as follows:

$$\varphi(x, \lambda) = \frac{\mu^2 x}{2} e^{\frac{1}{\mu^2 x}} \left\{ C_1 M_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 x} \right) + C_2 W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 x} \right) \right\}, \quad (3.6)$$

where C_1 and C_2 are arbitrary constants such that $C_1^2 + C_2^2 \neq 0$. The obvious next step is to “pin down” the two constants so as to make the obtained general $\varphi(x, \lambda)$ satisfy the boundary conditions (2.9) as well as make the pdf $q(x) \propto \mathfrak{m}(x) \varphi(x, \lambda)$ where $\mathfrak{m}(x)$ is given by (2.6) satisfy the normalization constraint (2.5). To that end, it follows from Polunchenko (2016) that C_1 must be set to zero to ensure the pdf $q(x)$ fulfills the entrance boundary condition (2.3). Hence, the general form (3.6) of $\varphi(x, \lambda)$ simplifies down to

$$\varphi(x, \lambda) = C \frac{\mu^2 x}{2} e^{\frac{1}{\mu^2 x}} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 x} \right), \quad (3.7)$$

where $C \neq 0$ is an arbitrary constant. Assuming for the moment that λ is known, the foregoing is sufficient to express $q(x)$ explicitly. Specifically, with the aid of (Gradshteyn and Ryzhik, 2007, Integral 7.623.7, p. 824) which states that

$$\begin{aligned} \int_1^{+\infty} (x-1)^{c-1} x^{a-c-1} e^{-\frac{zx}{2}} W_{a,b}(zx) dx &= \\ &= \Gamma(c) e^{-\frac{z}{2}} W_{a-c,b}(z), \text{ provided } \Re(c) > 0 \text{ and } \Re(z) > 0, \end{aligned} \quad (3.8)$$

where here and onward $\Gamma(z)$ denotes the Gamma function (see, e.g., Abramowitz and Stegun 1964, Chapter 6) one gets

$$\int_0^A \mathfrak{m}(x) \varphi(x, \lambda) dx = C e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right),$$

so that subsequently from (3.7) and the fact that $q(x) \propto \mathfrak{m}(x) \varphi(x, \lambda)$ where $\mathfrak{m}(x)$ is given by (2.6) it follows that the function

$$q(x) = \frac{\frac{1}{x} e^{-\frac{1}{\mu^2 x}} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 x} \right)}{e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right)} \mathbb{1}_{\{x \in [0, A]\}}$$

satisfies not only the entrance boundary condition (2.3) but also the normalization constraint (2.5).

It remains to find λ , i.e., the dominant eigenvalue of the operator \mathcal{D} given by (2.2). Formally, the problem is to find the largest root $\lambda \leq 0$ of the equation $\varphi(A, \lambda) = 0$ for a given $A > 0$. More concretely, in view of (2.4), (2.9) and (3.7), the equation to recover λ from is

$$W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right) = 0, \quad (3.9)$$

where $A > 0$ is fixed and $\xi(\lambda)$ is as in (3.4). This equation was previously analyzed by Polunchenko (2016), who, in particular, *explicitly* showed it to be consistent for any $A > 0$, just as predicted by the general Sturm–Liouville theory.

We have now established our first result, formally stated in the following theorem.

Theorem 3.1. *For any fixed detection threshold $A > 0$, the GSR statistic’s quasi-stationary density $q_A(x)$ is given by*

$$q_A(x) = \frac{\frac{1}{x} e^{-\frac{1}{\mu^2 x}} W_{1, \frac{\xi(\lambda)}{2}}\left(\frac{2}{\mu^2 x}\right)}{e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda)}{2}}\left(\frac{2}{\mu^2 A}\right)} \mathbb{1}_{\{x \in [0, A]\}}, \quad (3.10)$$

where λ is the largest (nonpositive) solution of the (surely consistent) equation

$$W_{1, \frac{\xi(\lambda)}{2}}\left(\frac{2}{\mu^2 A}\right) = 0$$

with $\xi(\lambda)$ as in (3.4). The respective quasi-stationary cdf $Q_A(x)$ is given by

$$Q_A(x) = \begin{cases} 1, & \text{for } x \geq A; \\ \frac{e^{-\frac{1}{\mu^2 x}} W_{0, \frac{\xi(\lambda)}{2}}\left(\frac{2}{\mu^2 x}\right)}{e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda)}{2}}\left(\frac{2}{\mu^2 A}\right)}, & \text{for } x \in [0, A]; \\ 0, & \text{otherwise.} \end{cases} \quad (3.11)$$

Proof. We only need to show (3.11). To that end, it suffices to integrate (3.10) with respect to x and evaluate the integral explicitly using the definite integral identity (3.8). \square

As complicated as the obtained formulae (3.10) and (3.11) may seem, both are all perfectly amenable to numerical evaluation, especially using *Mathematica*, which, as was previously demonstrated, e.g., by Polunchenko (2016) and by Linetsky (2004, 2007), can efficiently handle a variety of special functions, including the Whittaker W function. The only “delicate” part is the computation of the dominant eigenvalue λ , which, as can be seen from (3.10) and (3.11), is required for both the quasi-stationary pdf $q_A(x)$ as well as the corresponding cdf $Q_A(x)$. To that end, the equation (3.9) that determines λ can be easily treated numerically in *Mathematica* as well: one can get λ computed to within any desirable accuracy. By way of illustration of this point, we were able to get *Mathematica* compute λ to within as much as 400 (four hundred!) decimal places in about a minute, for each given $A > 0$, using an average laptop. Since 400 decimal places is extremely accurate, the corresponding numerically computed values of λ can be considered “exact”, and we shall use this terminology throughout the remainder of the paper. Figures 1 were obtained with the aid of our *Mathematica* script, and depict the quasi-stationary pdf $q_A(x)$, shown in Figure 1(a), and the corresponding cdf $Q_A(x)$, shown in Figure 1(b), as functions of $x \in [0, A]$, assuming $A = 20$ and μ is either $1/2$, 1 , or $3/2$. Note that, in either case, the pdf appears to be unimodal. As a matter of fact, this property of the quasi-stationary distribution will be formally established in Theorem 3.3 below.

On the other hand, the “explicitness” of formulae (3.10) and (3.11) allows to study the quasi-stationary distribution in an *analytic* manner, as we demonstrate in a series of results established next.

Let us begin with the following auxiliary result.

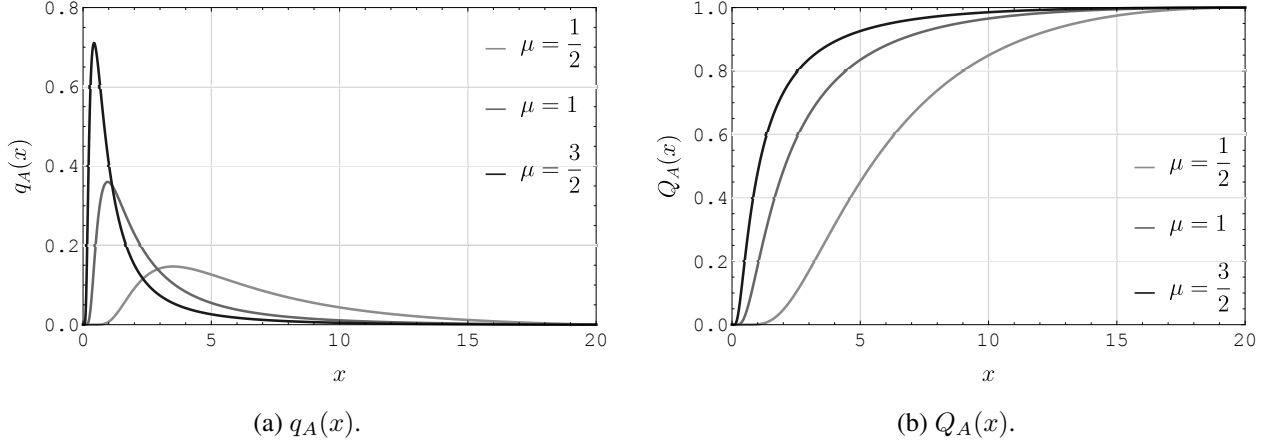


Figure 1. Quasi-stationary distribution's pdf $q_A(x)$ and cdf $Q_A(x)$ for $A = 20$ and $\mu = \{1/2, 1, 3/2\}$.

Lemma 3.1. *For any fixed $A > 0$ it is true that:*

(a) $\lim_{x \rightarrow 0+} q_A(x) = 0;$

(b) $\lim_{x \rightarrow 0+} \left[\frac{\partial}{\partial x} q_A(x) \right] = 0;$

(c) $A^2 \frac{\mu^2}{2} \left[\frac{\partial}{\partial x} q_A(x) \right] \Big|_{x=A} = \lambda.$

Proof. Part (a) can be easily deduced from (3.10) and

$$W_{1,b}(u) = u e^{-\frac{u}{2}} \left[1 + \mathcal{O}\left(\frac{1}{u}\right) \right] \text{ as } u \rightarrow +\infty, \text{ for any } b \in \mathbb{C},$$

which is a special case of the more general asymptotic property of the Whittaker W function

$$W_{a,b}(z) = z^a e^{-\frac{z}{2}} \left[1 + \mathcal{O}\left(\frac{1}{z}\right) \right] \text{ as } |z| \rightarrow +\infty, \text{ for any } b \in \mathbb{C}, \quad (3.12)$$

established, e.g., in (Whittaker and Watson, 1927, Section 16.3).

To prove part (b) explicitly, note that due to (Slater, 1960, Formula (2.4.18), p. 25) which states that

$$\frac{\partial^n}{\partial z^n} \left[z^{a+n-1} e^{-\frac{z}{2}} W_{a,b}(z) \right] = (-1)^n z^{a-1} e^{-\frac{z}{2}} W_{a+n,b}(z),$$

we have

$$\frac{\partial}{\partial z} \left[z e^{-\frac{z}{2}} W_{1,b}(z) \right] = -e^{-\frac{z}{2}} W_{2,b}(z),$$

so that by direct differentiation of (3.10) with respect to x one can subsequently see that

$$\frac{\partial}{\partial x} q_A(x) \propto \frac{1}{x^2} e^{-\frac{1}{\mu^2 x}} W_{2, \frac{\xi}{2}} \left(\frac{2}{\mu^2 x} \right), \quad (3.13)$$

and this goes to zero as $x \rightarrow 0+$ because

$$W_{2,b}(u) = u^2 e^{-\frac{u}{2}} \left[1 + \mathcal{O}\left(\frac{1}{u}\right) \right] \text{ as } u \rightarrow +\infty, \text{ for any } b \in \mathbb{C},$$

which is another special case of (3.12).

Finally, the assertion of part (c) will become apparent if one first integrates the master equation (2.1) through with respect to x over $[0, A]$, and then invokes parts (a) and (b) of this lemma, the normalization constraint (2.5), and the absorbing boundary condition (2.4) to simplify the result of the integration. \square

Lemma 3.1 is instrumental in getting the moments of the quasi-stationary distribution.

Theorem 3.2. *Let $A > 0$ be fixed, and suppose that Z is a Q_A -distributed random variable. Then the moments $M_n \triangleq \mathbb{E}[Z^n]$, $n \geq 0$, satisfy the recurrence*

$$M_n \left[\frac{\mu^2}{2} n(n-1) - \lambda \right] + n M_{n-1} = -\lambda A^n, \quad n \geq 1, \quad (3.14)$$

with $M_0 \equiv 1$.

Proof. It suffices to multiply the master equation (2.1) through by x^n and integrate the result with respect to x over the interval $[0, A]$ using integration by parts in conjunction with Lemma 3.1. \square

The recurrence (3.14) is essentially a nonhomogeneous first-order difference equation, and although it *can* be solved explicitly, the solution is too cumbersome to be of significant practical value. This notwithstanding, it is a simple exercise to infer from Theorem 3.2 that, if Z is a Q_A -distributed random variable and $A > 0$ is fixed, then $\mathbb{E}[Z] \triangleq M_1 = A + 1/\lambda$, and

$$\text{Var}[Z] = M_2 - M_1^2 = -\frac{\mu^2(A + 1/\lambda)^2 + 1/\lambda}{\mu^2 - \lambda},$$

where $\lambda \equiv \lambda_A$ is the largest (nonpositive) zero of equation (3.9). It is noteworthy that the foregoing formula trivially yields the following double inequality:

$$-\frac{1}{A} - \frac{1 + \sqrt{4\mu^2 A + 1}}{2\mu^2 A^2} \leq \lambda_A \leq -\frac{1}{A} - \frac{1 - \sqrt{4\mu^2 A + 1}}{2\mu^2 A^2} (< 0),$$

which makes it direct to see that $\lambda_A = -1/A + \mathcal{O}(A^{-3/2})$. Although the obtained bounds allow one to estimate λ_A for any given $A > 0$, and the accuracy improves as A gets higher, in the next subsection we will offer a more accurate (viz. order-three) large- A approximation to λ_A .

Another interesting property of the quasi-stationary distribution is its unimodality, which we already observed in Figures 1 above.

Theorem 3.3. *For any fixed detection threshold $A > 0$, the GSR statistic's quasi-stationary distribution given by (3.10) and by (3.11) is unimodal.*

Proof. The problem is effectively to show that for any given $A > 0$ the equation

$$\frac{d}{dx}q(x) = 0 \quad \text{with } x \in (0, A) \quad (3.15)$$

has only one solution, say $\tilde{x} \triangleq \tilde{x}_A$, and that this solution is such that the concavity condition

$$\left[\frac{d^2}{dx^2}q(x) \right] \Big|_{x=\tilde{x}} < 0 \quad (3.16)$$

is satisfied. We point out that by part (b) of Lemma 3.1 the point $x = 0$ is in the solution set of equation (3.15). However, it cannot be a maximum of $q(x)$, for $\lim_{x \rightarrow 0+} q(x) = 0$, as was asserted by part (a) of Lemma 3.1. Hence, the interval in (3.15) does not include the left end-point. The right end-point is not included because $q(A) = 0$ by design.

Let us first show that the concavity condition (3.16) is satisfied automatically for *any* solution $\tilde{x} \in (0, A)$ of equation (3.15). To that end, note that the master equation (2.1) considered at any such \tilde{x} can be rewritten in the form

$$\mu^2 \tilde{x}^2 \left[\frac{d^2}{dx^2}q(x) \right] \Big|_{x=\tilde{x}} = 2(\lambda - \mu^2)q(\tilde{x}),$$

whence, in view of Lemmas 2.1 and 2.2, the desired conclusion readily follows. Put another way, whatever extrema \tilde{x} the function $q(x)$ may have inside the interval $(0, A)$, they must all be the function's maxima.

It remains to show that $q(x)$ has exactly one extremum inside the interval $(0, A)$. To that end, it can be gleaned from (3.13) that the extrema of the function $q(x)$ are the solutions of the equation

$$W_{2, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 x} \right) = 0, \quad \text{such that } x \in (0, A),$$

where $\xi(\lambda)$ is given by (3.4), and $\lambda \leq 0$ is the largest solution of equation (3.9). At a slightly more abstract level, the question effectively is whether the equation $W_{2,b}(z) = 0$ with b fixed has exactly one (finite) solution, say \tilde{z} , that is to the right of the largest (but finite) positive zero of the function $W_{1,b}(z)$. This equation can be elegantly answered in the affirmative with the aid of the work of Milne (1914); see also (Sharma, 1938, Theorem V, p. 504). Specifically, Milne (1914) showed that if z_1 and z_2 are any two (finite) successive zeros of the function $W_{k,b}(z)$ with the indices k and b fixed, then the function $W_{k+1,b}(z)$ has exactly one zero located between z_1 and z_2 . With regard to the assumptions that k and b are to satisfy in order for the claim to hold true, while they are not explicitly stated in the paper, it is clear from Milne's (1914) proof that it certainly "goes through" if k is purely real and b is either purely real or purely imaginary. This is sufficient for our purposes, for in our case $k = 1$ and b is either purely real or purely imaginary, and the two zeros of $W_{1,b}(z)$ are $z_1 = 2/(\mu^2 A)$ and $z_2 = \infty$. Hence, the equation $W_{2,b}(z) = 0$ must have exactly one (finite) solution in between, and the proof is complete. \square

We remark that the stationary distribution $h(x) \triangleq \lim_{A \rightarrow +\infty} q_A(x) = \mathfrak{m}(x)$ supported on $x \in [0, +\infty)$, and where $\mathfrak{m}(x)$ is given by (2.6), is also unimodal: the mode is at $x^* = 1/\mu^2$. The mode of the quasi-stationary distribution is more difficult to obtain. However, if A is sufficiently large, then the asymptotic approximation of $q_A(x)$ offered in the next subsection can be used to get the mode approximately.

3.2. Asymptotic Analysis

We begin with two auxiliary results.

Lemma 3.2. *Let $A > 0$ be fixed, and suppose that $\{\lambda, \varphi(x, \lambda)\}$ is any eigenvalue-eigenfunction pair of the operator \mathcal{G} given by (2.8); recall that $\varphi(x, \lambda)$ is given by (3.7). Then*

$$\|\varphi(\cdot, \lambda)\|^2 = C^2 \left[\frac{\partial}{\partial \lambda} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right) \right] \left[\frac{\partial}{\partial u} W_{1, \frac{\xi(\lambda)}{2}}(u) \right] \Big|_{u=\frac{2}{\mu^2 A}}, \quad (3.17)$$

where constant $C \neq 0$ is the same constant as in the right-hand side of (3.7), and $\|\varphi(\cdot, \lambda)\|^2$ and $\xi(\lambda)$ are given by (2.10) and by (3.4), respectively.

Proof. See Polunchenko (2016). □

The foregoing lemma enables us to establish the following important monotonicity property of the spectrum of the operator \mathcal{D} given by (2.2) with respect to the detection threshold A .

Lemma 3.3. *The dominant eigenvalue $\lambda \equiv \lambda_A$ of the operator \mathcal{D} is a monotonically increasing function of $A > 0$. More concretely,*

$$\frac{d}{dA} \lambda_A > 0 \text{ for all } A > 0.$$

Proof. Let us show that the desired result is actually valid for *any* eigenvalue of the operator \mathcal{D} . Specifically, suppose temporarily that λ and $\varphi(x, \lambda)$ are an *arbitrary* eigenvalue-eigenfunction pair of \mathcal{G} . By implicit differentiation it is then direct to see from (3.9) that

$$\frac{d}{dA} \lambda_A = - \frac{\frac{\partial}{\partial A} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right)}{\frac{\partial}{\partial \lambda} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right)} = \frac{\frac{2}{\mu^2 A^2} \left[\frac{\partial}{\partial u} W_{1, \frac{\xi(\lambda)}{2}}(u) \right] \Big|_{u=\frac{2}{\mu^2 A}}}{\frac{\partial}{\partial \lambda} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right)},$$

and because from (3.17) we also have

$$\frac{\partial}{\partial \lambda} W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 A} \right) = \|\varphi(\cdot, \lambda)\|^2 / \left(C^2 \left[\frac{\partial}{\partial u} W_{1, \frac{\xi(\lambda)}{2}}(u) \right] \Big|_{u=\frac{2}{\mu^2 A}} \right),$$

one can readily conclude that

$$\frac{d}{dA} \lambda(A) = \frac{2}{\mu^2 A^2} \left(C \left[\frac{\partial}{\partial u} W_{1, \frac{\xi(\lambda)}{2}}(u) \right] \Big|_{u=\frac{2}{\mu^2 A}} \right)^2 / \|\varphi(\cdot, \lambda)\|^2,$$

which gives the desired (more general) result. □

Lemma 3.3 explicitly shows that to say “large A ” is the same as to say “small $\lambda \equiv \lambda_A$ ”. This means, in particular, that $\lim_{A \rightarrow +\infty} \lambda_A = 0$, which makes perfect sense. The large- A behavior of the quasi-stationary distribution can therefore be extracted from the Taylor series of the function $W_{1, \frac{\xi(\lambda)}{2}}(x)$ present in the numerator of the fraction in the right-hand side of (3.10) expanded with respect to λ around zero, assuming that $x \in [0, A]$ is fixed. Specifically, noting from (3.4) that $\xi(0) = 1$, the problem is to obtain the expansion

$$W_{1, \frac{\xi(\lambda)}{2}}(x) = W_{1, \frac{1}{2}}(x) + \sum_{k=1}^n \left[\frac{\partial^k}{\partial \lambda^k} W_{1, \frac{\xi(\lambda)}{2}}(x) \right] \Big|_{\lambda=0} \frac{\lambda^k}{k!} + \mathcal{O}(\lambda^{n+1}), \quad (3.18)$$

up to a desired order $n \geq 1$; while the order n can, in principle, be 0, shrinking the foregoing Taylor series to just the first term would not yield a particularly accurate approximation. With regard to the denominator of the fraction in the right-hand of (3.10), it can be safely forgotten about, because

$$\lim_{A \rightarrow +\infty} e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda_A)}{2}} \left(\frac{2}{\mu^2 A} \right) = 1,$$

which follows from the fact that $\lim_{x \rightarrow 0+} x^x = 1$ and the Whittaker W function’s small-argument asymptotics

$$W_{0,b}(u) \sim \frac{\Gamma(2b)}{\Gamma(b+1/2)} u^{\frac{1}{2}-b} e^{-\frac{u}{2}} \text{ as } u \rightarrow 0+, \text{ for any } b \in \mathbb{C},$$

which, in turn, is a special case of

$$W_{a,b}(u) \sim \frac{\Gamma(2b)}{\Gamma(b-a+1/2)} u^{\frac{1}{2}-b} e^{-\frac{u}{2}} \text{ as } u \rightarrow 0+, \text{ for any } a, b \in \mathbb{C}.$$

The first term of the sought series (3.18) can be worked out from (Buchholz, 1969, Identity (28a), p. 23), according to which

$$W_{a, a-\frac{1}{2}}(z) = z^a e^{-\frac{z}{2}},$$

so that

$$W_{1, \frac{1}{2}}(z) = z e^{-\frac{z}{2}},$$

and we can proceed to finding the higher-order terms. To that end, the task requires the evaluation of the partial derivatives

$$\left[\frac{\partial^k}{\partial b^k} W_{1,b}(x) \right] \Big|_{b=\frac{1}{2}},$$

for a fixed $x \in [0, A]$ and for all k ’s starting from 1 and up to the desired order $n \geq 1$ of the expansion (3.18). We shall now set $n = 3$ and find the necessary three derivatives, so as to then obtain the corresponding *third-order* expansion. This will yield a sufficiently accurate approximation of the quasi-stationary distribution, and augmenting the expansion with higher-order terms—though *is* possible—would be superfluous.

To proceed, let us recall two special functions that will be needed below. The first function we will need is the exponential integral

$$\text{Ei}(x) \triangleq \begin{cases} -\int_{-x}^{\infty} \frac{e^{-y}}{y} dy, & \text{for } x < 0; \\ -\lim_{\varepsilon \rightarrow +0} \left[\int_{-x}^{-\varepsilon} \frac{e^{-y}}{y} dy + \int_{\varepsilon}^{\infty} \frac{e^{-y}}{y} dy \right], & \text{for } x > 0, \end{cases}$$

whose basic properties are summarized, e.g., in (Abramowitz and Stegun, 1964, Chapter 5). More specifically, we will need the function $E_1(x) \triangleq -\text{Ei}(-x)$, but restricted to positive values of the argument, so that

$$E_1(x) \triangleq \int_x^\infty \frac{e^{-y}}{y} dy = x e^{-x} \int_0^{+\infty} e^{-xy} \log(1+y) dy, \quad x > 0, \quad (3.19)$$

where the second equality is due to (Gradshteyn and Ryzhik, 2007, Integral 4.337.1, p. 572) which is

$$\int_0^{+\infty} e^{-ay} \log(b+y) dy = \frac{1}{a} [\log b - e^{ab} \text{Ei}(-ba)], \quad |\arg b| < \pi, \quad \Re(a) > 0.$$

The second function we will need is the Meijer G -function, introduced in the seminal work of Meijer (1936). The Meijer G -function is defined as the Mellin-Barnes integral

$$G_{p,q}^{m,n} \left(\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| z \right) \triangleq \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\prod_{k=1}^m \Gamma(b_k - s) \prod_{j=1}^n \Gamma(1 - a_j + s)}{\prod_{k=m+1}^q \Gamma(1 - b_k + s) \prod_{j=n+1}^p \Gamma(a_j - s)} z^s ds, \quad (3.20)$$

where i denotes the imaginary unit, i.e., $i \triangleq \sqrt{-1}$, the integers m, n, p , and q are such that $0 \leq m \leq q$ and $0 \leq n \leq p$, and the contour \mathcal{C} is closed in an appropriate way to ensure the convergence of the integral. It is also required that no $a_j - b_k$ be an integer. The function is a very general function, and includes, as special cases, not only all elementary functions, but a number of special functions as well. An extensive list of special cases of the Meijer G -function can be found, e.g., in the classical special functions handbook of Prudnikov et al. (1990), which also includes a summary of the function's basic properties. We will need the following particular case of the Meijer G -function:

$$G_{2,3}^{3,1} \left(\begin{matrix} 0, 1 \\ 0, 0, 0 \end{matrix} \middle| x \right) = \int_x^{+\infty} e^y E_1(y) \frac{dy}{y} = \int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y}, \quad x > 0; \quad (3.21)$$

see Appendix A for a proof.

We are now in a position to formulate the main result that will subsequently enable us to obtain the expansion (3.18) explicitly, up to the third order.

Lemma 3.4. *For any fixed $x \geq 0$ the following identities hold true:*

- (a) $\left[\frac{\partial}{\partial b} W_{1,b}(x) \right] \Big|_{b=\frac{1}{2}} = e^{-\frac{x}{2}};$
- (b) $\left[\frac{\partial^2}{\partial b^2} W_{1,b}(x) \right] \Big|_{b=\frac{1}{2}} = 2 e^{-\frac{x}{2}} \left\{ e^x E_1(x) + x G_{2,3}^{3,1} \left(\begin{matrix} 0, 1 \\ 0, 0, 0 \end{matrix} \middle| x \right) \right\};$
- (c) $\left[\frac{\partial^3}{\partial b^3} W_{1,b}(x) \right] \Big|_{b=\frac{1}{2}} = 6 e^{-\frac{x}{2}} G_{2,3}^{3,1} \left(\begin{matrix} 0, 1 \\ 0, 0, 0 \end{matrix} \middle| x \right).$

Our proof of this lemma is offered in Appendix B, and it is based on repetitive differentiation of the Mellin integral transform (see, e.g., Oberhettinger 1974, pp. 1–10) of the function $f(x) \triangleq e^{-\frac{x}{2}} W_{1,b}(x)$, $x \geq 0$, with respect to the second index b . As an aside, we note that, apparently, the identities established in Lemma 3.4 have not yet been derived in the theory of special functions (at least we were unable to find an existing reference for any one of them).

It is now straightforward to obtain the sought third-order expansion. Specifically, Lemma 3.4 and the basic differentiation chain rule together lead to the following result.

Theorem 3.4. For any $x \geq 0$ it holds true that

$$W_{1, \frac{\xi(\lambda)}{2}} \left(\frac{2}{\mu^2 x} \right) = \frac{2}{\mu^2} e^{-\frac{1}{\mu^2 x}} \left\{ \frac{1}{x} + \lambda + \frac{2}{\mu^2} L \left(\frac{2}{\mu^2 x} \right) \lambda^2 + \left(\frac{2}{\mu^2} \right)^2 \left[G_{2,3}^{3,1} \left(\frac{0,1}{0,0,0} \middle| \frac{2}{\mu^2 x} \right) - 2L \left(\frac{2}{\mu^2 x} \right) \right] \lambda^3 \right\} + \mathcal{O}(\lambda^4), \quad (3.22)$$

where $\xi(\lambda)$ is as in (3.4), and

$$L(x) \triangleq e^x E_1(x) - 1 + x G_{2,3}^{3,1} \left(\frac{0,1}{0,0,0} \middle| x \right). \quad (3.23)$$

As the first application of the foregoing theorem, let us obtain an order-one, order-two, and order-three approximations of the dominant eigenvalue λ . To that end, it suffices to apply (3.22) to approximate the Whittaker W function sitting in the left-hand side of equation (3.9). Specifically, in view of (3.22), approximated up to the first order, the equation is simply $1/A + \lambda = 0$, so that

$$\lambda^* \equiv \lambda_A^* \triangleq -\frac{1}{A}$$

is λ 's first-order approximation.

Likewise, getting λ 's second-order approximation is a matter of solving the quadratic equation:

$$\frac{2}{\mu^2} L \left(\frac{2}{\mu^2 A} \right) \lambda^2 + \lambda + \frac{1}{A} = 0,$$

which may or may not have *real* solutions, depending on whether A is large enough. Assuming it is, the quadratic equation has two distinct real roots, both negative, and the one closest to zero should be used as λ 's second-order approximation. Written explicitly, the second-order approximation is as follows:

$$\lambda^{**} \equiv \lambda_A^{**} \triangleq -\frac{\mu^2}{4} \frac{1 - \sqrt{1 - \frac{8}{\mu^2 A} L \left(\frac{2}{\mu^2 A} \right)}}{L \left(\frac{2}{\mu^2 A} \right)},$$

and we reiterate that it requires the detection threshold $A > 0$ to be sufficiently large.

Finally, the third-order approximation is determined by the cubic equation:

$$\frac{1}{A} + \lambda + \frac{2}{\mu^2} L \left(\frac{2}{\mu^2 A} \right) \lambda^2 + \left(\frac{2}{\mu^2} \right)^2 \left[G_{2,3}^{3,1} \left(\frac{0,1}{0,0,0} \middle| \frac{2}{\mu^2 A} \right) - 2L \left(\frac{2}{\mu^2 A} \right) \right] \lambda^3 = 0,$$

which, again for sufficiently large A 's, has exactly one real root, because the coefficient in front of λ^3 is negative. It is that single real solution that, when exists, should be used as λ 's third-order approximation, i.e., as $\lambda^{***} \equiv \lambda_A^{***}$. While it is possible to express λ^{***} explicitly, the formula is simply too “bulky” to state, and for this reason only we shall not present it.

The quality of each of the three approximations of λ can be judged from Table 1, which reports the “exact” (i.e., computed to within 400 decimal places of accuracy) value of λ , and the corresponding three approximations thereof for a handful of values of $A > 0$, and assuming $\mu = 1$. Specifically, the table provides only the first 12 decimal places

Table 1. Dominant eigenvalue λ and its first-, second- and, third-order approximations $-\lambda^*$, $-\lambda^{**}$, $-\lambda^{***}$ as functions of A for $\mu = 1$.

A	$-\lambda$	$-\lambda^*$	$-\lambda^{**}$	$-\lambda^{***}$
20	0.058856148622	0.05	0.059819055496	0.058817735494
30	0.037786534271	0.0333333333333	0.03811217223	0.03777661428
40	0.027727324417	0.025	0.027880519395	0.027723505394
50	0.02186160095	0.02	0.021947421685	0.02185977578
100	0.010563106075	0.01	0.010577520296	0.010562921283
500	0.002033066472	0.002	0.002033295282	0.002033065611
1,000	0.0010095172	0.001	0.001009554734	0.001009517118
10,000	0.000100139278	0.0001	0.000100139359	0.000100139278

of $-\lambda$, $-\lambda^*$, $-\lambda^{**}$, and of $-\lambda^{***}$, and the negation is done sheerly for convenience. It is evident from the table that the first-order approximation λ^* performs somewhat poorly, unless the detection threshold is on the order of thousands, which is considered high in practice. The second-order approximation λ^{**} performs much better, even when the detection threshold is moderate. The third-order approximation λ^{***} is the most accurate of the three approximations, and is fairly close to the “exact” value of λ for A ’s on the order of tens, which is rather low from a practical standpoint. However, in terms of computational convenience, the ranking of the the three approximations is the exact opposite of their accuracy ranking.

We are now in a position to offer an order-one, order-two, and order-three “large- A ” approximations of the quasi-stationary pdf $q_A(x)$ given by (3.10). Let the approximations be denoted as $q_A^*(x)$, $q_A^{**}(x)$, and $q_A^{***}(x)$, respectively, which is in analogy to the notation adapted for the three approximations of λ . By combining Theorem 3.4 and (3.10)

we obtain

$$\begin{aligned}
q_A^*(x) &= \frac{\frac{2}{\mu^2 x} e^{-\frac{2}{\mu^2 x}}}{e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda^*)}{2}} \left(\frac{2}{\mu^2 A} \right)} \left\{ \frac{1}{x} + \lambda^* \right\} \mathbb{1}_{\{x \in [0, A]\}}; \\
q_A^{**}(x) &= \frac{\frac{2}{\mu^2 x} e^{-\frac{2}{\mu^2 x}}}{e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda^{**})}{2}} \left(\frac{2}{\mu^2 A} \right)} \left\{ \frac{1}{x} + \lambda^{**} + \frac{2}{\mu^2} L \left(\frac{2}{\mu^2 x} \right) (\lambda^{**})^2 \right\} \mathbb{1}_{\{x \in [0, A]\}}; \\
q_A^{***}(x) &= \frac{\frac{2}{\mu^2 x} e^{-\frac{2}{\mu^2 x}}}{e^{-\frac{1}{\mu^2 A}} W_{0, \frac{\xi(\lambda^{***})}{2}} \left(\frac{2}{\mu^2 A} \right)} \times \\
&\quad \times \left\{ \frac{1}{x} + \lambda^{***} + \frac{2}{\mu^2} L \left(\frac{2}{\mu^2 x} \right) (\lambda^{***})^2 + \right. \\
&\quad \left. + \left(\frac{2}{\mu^2} \right)^2 \left[G_{2,3}^{3,1} \left(\frac{0,1}{0,0,0} \middle| \frac{2}{\mu^2 x} \right) - 2L \left(\frac{2}{\mu^2 x} \right) \right] (\lambda^{***})^3 \right\} \mathbb{1}_{\{x \in [0, A]\}},
\end{aligned}$$

where $L(x)$ is defined in (3.23), and the corresponding three approximations of λ are computed as described above.

To get an idea as to the accuracy of the obtained approximations, let us look at Figures 2. Specifically, Figure 2(a) shows the “exact” pdf $q_A(x)$ and the three approximations thereof, all as functions of $x \in [0, A]$, assuming $A = 20$ and $\mu = 1$. Figure 2(b) shows the corresponding absolute errors, i.e., the quantities $|q_A(x) - q_A^*(x)|$, $|q_A(x) - q_A^{**}(x)|$, and $|q_A(x) - q_A^{***}(x)|$. The detection threshold is intentionally set so low, for otherwise the three approximations would be closer to the actual pdf, and the corresponding errors would be harder to notice. Observe from the figures that the first-order approximation is noticeably off. Recalling the numbers reported earlier in Table 1, this is a direct consequence of the threshold being set so low. However, in spite of the low threshold, the second- and third-order approximations, whose corresponding curves appear to nearly coincide in the figures, are both fairly close to the actual “exact” pdf, across the entire range of values of $x \in [0, A]$. While the quality of all three approximations improves as the threshold gets bigger, the second- and third-order approximations each become practically indistinguishable from the exact pdf as soon as the threshold is in the hundreds, which is about the range often used in practice.

We conclude this subsection with a remark concerning the function $L(x)$ introduced in (3.23). This function plays an important role in the minimax theory of quickest change-point detection, since it is the key ingredient of the universal lowerbound (obtained, e.g., in Feinberg and Shiryaev 2006, Lemma 2.2) on the still-unknown optimal value of the detection delay penalty-function introduced by Pollak (1985); for the discrete-time setting of the problem, the equivalent of this lowerbound was introduced by Moustakides et al. (2011). Therefore, function $L(x)$ is essential in assessing the efficiency of a detection procedure, and, in particular, possibly proving that the procedure of interest is either exactly or nearly Pollak-minimax. Precisely this approach was employed by Burnaev et al. (2009) to prove that the randomized version of the (Generalized) Shiryaev–Roberts procedure is asymptotically order-two Pollak-minimax. The main apparatus used throughout our investigation, i.e., the Whittaker W function and the calculus associated with it, may help improve the second-order minimaxity established by Burnaev et al. (2009). To that end, observe that

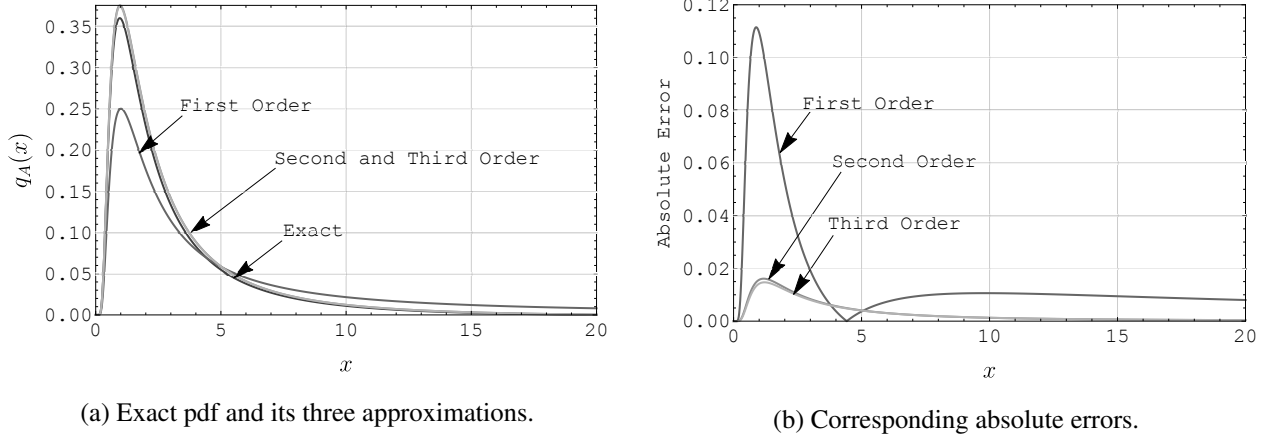


Figure 2. Quasi-stationary distribution's pdf $q_A(x)$ and its three approximations $q_A^*(x)$, $q_A^{**}(x)$, and $q_A^{***}(x)$ as functions of $x \in [0, A]$ for $A = 20$ and $\mu = 1$.

part (c) of Lemma 3.4 provides a link between function $L(x)$ and the Whittaker W function. Consequently, the lowerbound, which is effectively defined in (3.23), may be reexpressed as the second derivative of the corresponding Whittaker function (the derivative is with respect to the second index). This alternative form of the lowerbound might offer a new insight into the problem, and a further investigation of this appears to be worthwhile.

4. CONCLUDING REMARKS

The focus of this work was on the quasi-stationary distribution of the (Generalized) Shiryayev–Roberts (GSR) diffusion from quickest change-point detection theory. We obtained exact formulae for the distribution's cdf and pdf; the formulae are valid for any threshold that the GSR diffusion is stopped at. We also found the distribution's entire moment series, and proved the distribution to be unimodal for any threshold. More importantly, we offered an order-three large-threshold asymptotic approximation of the quasi-stationary distribution. The derivation of the approximation required establishing new identities on certain special functions. This work's findings are of direct application to quickest change-point detection, and we expect to make the transition in a subsequent paper.

A. PROOF OF FORMULA 3.21

We first note that the identity

$$\int_x^{+\infty} e^y E_1(y) \frac{dy}{y} = \int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y}, \quad x > 0,$$

where $E_1(x)$ is the exponential integral defined in (3.19), has already been established, e.g., by (Feinberg and Shiryayev, 2006, pp. 453–454). It therefore remains to show only that

$$G_{2,3}^{3,1} \left(\begin{matrix} 0,1 \\ 0,0,0 \end{matrix} \middle| x \right) = \int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y}, \quad x > 0,$$

where G denotes the Meijer G -function whose general definition is given by (3.20). To that end, the idea is to use (Prudnikov et al., 1990, Identity 8.4.6.5, p. 537) according to which

$$\log(1+x) = G_{2,2}^{1,2} \left(\begin{matrix} 1,1 \\ 1,0 \end{matrix} \middle| x \right),$$

so that

$$\int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y} = \int_0^{+\infty} e^{-xy} G_{2,2}^{1,2} \left(\begin{matrix} 1,1 \\ 1,0 \end{matrix} \middle| y \right) \frac{dy}{y},$$

whence, in view of the definite integral

$$\begin{aligned} \int_0^{+\infty} e^{-\beta x} G_{p,q}^{m,n} \left(\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| \alpha x \right) \frac{dx}{x^c} &= \beta^{c-1} G_{p+1,q}^{m,n+1} \left(\begin{matrix} c, a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| \frac{\alpha}{\beta} \right), \\ \text{provided } p+q &< 2(m+n), |\arg \alpha| < (2m+2n-p-q) \frac{\pi}{2}, \\ |\arg \beta| &< \frac{\pi}{2}, \text{ and } \Re(b_j - c) > -1, \quad j = 1, \dots, m, \end{aligned}$$

as given, e.g., by (Gradshteyn and Ryzhik, 2007, Identity 7.813.1, p. 853), it follows that

$$\int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y} = G_{3,2}^{1,3} \left(\begin{matrix} 1,1,1 \\ 1,0 \end{matrix} \middle| \frac{1}{x} \right),$$

and to end the proof, it suffices to recall one of the basic properties of the Meijer G -function, viz. one asserting that

$$G_{p,q}^{m,n} \left(\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| z \right) = G_{q,p}^{n,m} \left(\begin{matrix} 1-b_1, \dots, 1-b_q \\ 1-a_1, \dots, 1-a_p \end{matrix} \middle| \frac{1}{z} \right),$$

as given, e.g., by (Prudnikov et al., 1990, Identity 8.2.2.14, p. 521), and also observe that, by definition (3.20), any Meijer G -function is symmetric with respect to a_1, \dots, a_n , with respect to a_{n+1}, \dots, a_p , with respect to b_1, \dots, b_m , and with respect to b_{m+1}, \dots, b_q —each group of parameters treated separately.

B. PROOF OF LEMMA 3.4

The centerpiece of our proof is the identity

$$\int_0^{+\infty} x^{s-1} e^{-\frac{cx}{2}} W_{a,b}(cx) dx = \frac{\Gamma(1/2+b+s) \Gamma(1/2-b+s)}{c^s \Gamma(1-a+s)}, \quad s > -1/2 \pm \Re(b);$$

cf., e.g., (Oberhettinger, 1974, Identity 13.52, p. 146). This identity is the Mellin transform of the function $f(x) \triangleq e^{-\frac{cx}{2}} W_{a,b}(cx)$ with a, b , and c fixed, and $x \geq 0$. We are interested in the case when $c = a = 1$, and when b is either purely real and such that $0 \leq b \leq 1/2$, or purely imaginary (i.e., $\Re(b) = 0$). In this case the above identity reduces to

$$\int_0^{+\infty} x^{s-1} e^{-\frac{x}{2}} W_{1,b}(x) dx = \frac{\Gamma(1/2+b+s) \Gamma(1/2-b+s)}{\Gamma(s)}, \quad s > 0,$$

and the entire proof is based on successive differentiation of the foregoing with respect to b , followed by the evaluation of the result at $b = 1/2$.

To show part (a), observe first that a differentiation of the above integral identity through with respect to b gives

$$\begin{aligned} \int_0^{+\infty} x^{s-1} e^{-\frac{x}{2}} \left[\frac{\partial}{\partial b} W_{1,b}(x) \right] dx &= \frac{\Gamma(1/2+b+s) \Gamma(1/2-b+s)}{\Gamma(s)} \times \\ &\times [\psi_0(1/2+b+s) - \psi_0(1/2-b+s)], \quad s > 0, \end{aligned} \tag{B.1}$$

where here and onward $\psi_0(z)$ denotes the digamma function defined as $\psi_0(z) \triangleq d \log \Gamma(z)/dz$; for the basic background on the digamma function, see, e.g., (Abramowitz and Stegun, 1964, Section 6.3). The substitution $b = 1/2$ turns (B.1) into

$$\int_0^{+\infty} x^{s-1} e^{-\frac{x}{2}} \left[\frac{\partial}{\partial b} W_{1,b}(x) \right] \Big|_{b=1/2} dx = \Gamma(s), \quad s > 0,$$

because $\psi_0(z+1) - \psi_0(z) = 1/z$, as given, e.g., by (Abramowitz and Stegun, 1964, Identity 6.3.5, p. 258). Finally, since $\Gamma(s)$ for $s > 0$, is the Mellin transform of the function $f(x) \triangleq e^{-x}$, $x \geq 0$, one can deduce that

$$e^{-\frac{x}{2}} \left[\frac{\partial}{\partial b} W_{1,b}(x) \right] \Big|_{b=1/2} = e^{-x}, \quad x \geq 0,$$

whence the desired result.

Next, by differentiating (B.1) through with respect to b we obtain

$$\begin{aligned} \int_0^{+\infty} x^{s-1} e^{-\frac{x}{2}} \left[\frac{\partial^2}{\partial b^2} W_{1,b}(x) \right] dx &= \frac{\Gamma(1/2 + b + s) \Gamma(1/2 - b + s)}{\Gamma(s)} \times \\ &\times \left\{ [\psi_0(1/2 + b + s) - \psi_0(1/2 - b + s)]^2 + \right. \\ &\left. + [\psi_1(1/2 + b + s) + \psi_1(1/2 - b + s)] \right\}, \quad s > 0, \end{aligned} \quad (\text{B.2})$$

where here and onward $\psi_1(z) \triangleq d\psi_0(z)/dz$, i.e., $\psi_1(z)$ is the trigamma function, which is a particular case of the more general polygamma function $\psi_m(z) \triangleq d^m \psi_0(z)/dz^m$, $m = 0, 1, 2, \dots$; for the basic background on the polygamma function, see, e.g., (Abramowitz and Stegun, 1964, Section 6.4). The substitution $b = 1/2$ turns (B.2) into

$$\int_0^{+\infty} x^{s-1} e^{-\frac{x}{2}} \left[\frac{\partial^2}{\partial b^2} W_{1,b}(x) \right] \Big|_{b=1/2} dx = 2 s \Gamma(s) \psi_1(s), \quad s > 0, \quad (\text{B.3})$$

and the derivation exploits the recurrence $\psi_0(z+1) - \psi_0(z) = 1/z$ already used above, and the recurrence $\psi_1(z+1) - \psi_1(z) = -1/z^2$, which is a special case of (Abramowitz and Stegun, 1964, Identity 6.4.6, p. 260) stating that

$$\psi_m(z+1) - \psi_m(z) = (-1)^m \frac{m!}{z^{m+1}}, \quad m = 0, 1, 2, \dots \quad (\text{B.4})$$

To “undo” the Mellin transform (B.3), recall that for any two functions $g_1(x)$ and $g_2(x)$ whose Mellin transforms are $G_1(s)$ and $G_2(s)$, respectively, the Mellin transform of their *multiplicative* convolution

$$f(x) = \int_0^{+\infty} g_1\left(\frac{x}{t}\right) g_2(t) \frac{dt}{t}$$

is the product $G_1(s) G_2(s)$. See, e.g., (Oberhettinger, 1974, Identity 1.14, p. 12). Since by (Oberhettinger, 1974, Identity 4.13, p. 35) the Mellin transform of the function

$$g_2(x) = \frac{\log x}{x-1} \mathbb{1}_{\{0 < x < 1\}}$$

is precisely $\psi_1(s)$ with $s > 0$, i.e., the trigamma function, and because by (Oberhettinger, 1974, Identity 1.10, p. 12) the Mellin transform of the function $g_1(x) = 2x e^{-x}$, $x \geq 0$, is $2s \Gamma(s)$, also with $s > 0$, it follows that

$$\left[\frac{\partial^2}{\partial b^2} W_{1,b}(x) \right] \Big|_{b=1/2} = 2 e^{\frac{x}{2}} \int_0^1 \frac{x}{t} e^{-\frac{x}{t}} \frac{1}{t-1} \log t \frac{dt}{t}, \quad x \geq 0.$$

The change of integration variables from t to $y \equiv y(t) \triangleq 1/t - 1$ allows to rewrite the foregoing as follows:

$$\left[\frac{\partial^2}{\partial b^2} W_{1,b}(x) \right] \Big|_{b=1/2} = 2x e^{-x} \left\{ \int_0^{+\infty} e^{-xy} \log(1+y) dy + \int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y} \right\}, \quad x \geq 0,$$

which, in view of (3.19) and (3.21), can be seen to be precisely part (b).

Proving part (c) involves exactly the same steps. By differentiating (B.2) through with respect to b and then evaluating the result at $b = 1/2$ one obtains

$$\int_0^{+\infty} x^{s-1} e^{-\frac{x}{2}} \left[\frac{\partial^3}{\partial b^3} W_{1,b}(x) \right] \Big|_{b=1/2} dx = 6 \Gamma(s) \psi_1(s), \quad s > 0,$$

and, as before, the key role in the derivation is played by the recurrence (B.4) with m equal to 0, 1, and 2. This implies that

$$\left[\frac{\partial^3}{\partial b^3} W_{1,b}(x) \right] \Big|_{b=1/2} = 6 e^{\frac{x}{2}} \int_0^1 e^{-\frac{x}{t}} \frac{1}{t-1} \log t \frac{dt}{t}, \quad x \geq 0,$$

which, using the substitution $y \equiv y(t) \triangleq 1/t - 1$, can be rewritten as

$$\left[\frac{\partial^3}{\partial b^3} W_{1,b}(x) \right] \Big|_{b=1/2} = 6 e^{-\frac{x}{2}} \int_0^{+\infty} e^{-xy} \log(1+y) \frac{dy}{y}, \quad x \geq 0,$$

and this, in view of (3.21), is precisely part (c).

ACKNOWLEDGEMENT

The author is grateful to the Editor-in-Chief, Nitis Mukhopadhyay (University of Connecticut–Storrs) and to the anonymous referee for the constructive feedback provided on the first draft of the paper that helped improve the quality of the manuscript and shape its current form.

The author's effort was partially supported by the Simons Foundation via a Collaboration Grant in Mathematics under Award #304574.

REFERENCES

- Abramowitz, M. and Stegun, I. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th edition, Washington: United States Department of Commerce, National Bureau of Standards.
- Basseville, M. and Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Application*, Englewood Cliffs: Prentice Hall.
- Buchholz, H. (1969). *The Confluent Hypergeometric Function*, New York: Springer. Translated from German into English by H. Lichtblau and K. Wetzel.

- Burnaev, E. V. (2009). On a Nonrandomized Change-point Detection Method Second-Order Optimal in the Minimax Brownian Motion Problem, in *Proceedings of X All-Russia Symposium on Applied and Industrial Mathematics (Fall open session)*, October 1–8, Sochi, Russia (in Russian).
- Burnaev, E. V., Feinberg, E. A., and Shiryaev, A. N. (2009). On Asymptotic Optimality of the Second Order in the Minimax Quickest Detection Problem of Drift Change for Brownian Motion, *Theory of Probability and Its Applications* 53: 519–536.
- Cattiaux, P., Collet, P., Lambert, A., Martínez, S., Méléard, S., and Martín, J. S. (2009). Quasi-stationary Distributions and Diffusion Models in Population Dynamics, *Annals of Probability* 37: 1926–1969.
- Coddington, E. A. and Levinson, N. (1955). *Theory of Ordinary Differential Equations*, New York: McGraw-Hill.
- Dunford, N. and Schwartz, J. T. (1963). *Linear Operators. Part II: Spectral Theory. Self Adjoint Operators in Hilbert Space*, New York: Wiley.
- Feinberg, E. A. and Shiryaev, A. N. (2006). Quickest Detection of Drift Change for Brownian Motion in Generalized Bayesian and Minimax Settings, *Statistics & Decisions* 24: 445–470.
- Feller, W. (1952). The Parabolic Differential Equations and the Associated Semi-groups of Transformations, *Annals of Mathematics* 55: 468–519.
- Fulton, C. T., Pruess, S., and Xie, Y. (1999). *The Automatic Classification of Sturm–Liouville Problems*. Technical report, Florida Institute of Technology. Available online at: <http://citeseerx.ist.psu.edu/viewdoc/versions?doi=10.1.1.50.7591>.
- Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products*, 7th edition, New York: Academic Press.
- Itô, K. and McKean, Jr., H. P. (1974). *Diffusion Processes and Their Sample Paths*, Berlin: Springer.
- Kent, J. T. (1980). Eigenvalue Expansions for Diffusion Hitting Times, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 52: 309–319.
- Levitan, B. M. (1950). *Eigenfunction Expansions of Second-Order Differential Equations*, Leningrad: Gostechizdat (in Russian).
- Levitan, B. M. and Sargsjan, I. S. (1975). *Introduction to Spectral Theory: Selfadjoint Ordinary Differential Operators*, Providence: American Mathematical Society.
- Linetsky, V. (2004). Spectral Expansions for Asian (Average Price) Options, *Operations Research* 52: 856–867.
- Linetsky, V. (2007). *Spectral Methods in Derivative Pricing*, Netherlands: North–Holland.
- Mandl, P. (1961). Spectral Theory of Semi-groups Connected with Diffusion Processes and Its Application, *Czechoslovak Mathematical Journal* 11: 558–569.
- Meijer, C. S. (1936). Über Whittakersche bzw. Besselsche Funktionen und deren Produkte, *Nieuw Archief voor Wiskunde, Serie 2* 18: 10–39 (in German).
- Milne, A. (1914). On the Roots of the Confluent Hypergeometric Functions, in *Proceedings of Edinburgh Mathematical Society* 33: 48–64.

- Moustakides, G. V. (2004). Optimality of the CUSUM Procedure in Continuous Time, *Annals of Statistics* 32: 302–315.
- Moustakides, G. V., Polunchenko, A. S., and Tartakovsky, A. G. (2011). A Numerical Approach to Performance Analysis of Quickest Change-point Detection Procedures, *Statistica Sinica* 21: 571–596.
- Oberhettinger, F. (1974). *Tables of Mellin Transforms*, New York: Springer.
- Pollak, M. (1985). Optimal Detection of a Change in Distribution, *Annals of Statistics* 13: 206–227.
- Pollak, M. and Siegmund, D. (1985). A Diffusion Process and Its Applications to Detecting a Change in the Drift of Brownian Motion, *Biometrika* 72: 267–280.
- Pollak, M. and Siegmund, D. (1986). Convergence of Quasi-stationary to Stationary Distributions for Stochastically Monotone Markov Processes, *Journal of Applied Probability* 23: 215–220.
- Polunchenko, A. S. (2016). Exact Distribution of the Generalized Shiryaev–Roberts Stopping Time Under the Minimax Brownian Motion Setup, *Sequential Analysis* 35: 108–143.
- Polunchenko, A. S. and Sokolov, G. (2016). An Analytic Expression for the Distribution of the Generalized Shiryaev–Roberts Diffusion, *Methodology and Computing in Applied Probability*, in press. Available online at: <http://link.springer.com/article/10.1007/s11009-016-9478-7>.
- Polunchenko, A. S. and Tartakovsky, A. G. (2010). On Optimality of the Shiryaev–Roberts Procedure for Detecting a Change in Distribution, *Annals of Statistics* 38: 3445–3457.
- Polunchenko, A. S. and Tartakovsky, A. G. (2012). State-of-the-Art in Sequential Change-point Detection, *Methodology and Computing in Applied Probability* 14: 649–684.
- Poor, H. V. and Hadjiladis, O. (2009). *Quickest Detection*, New York: Cambridge University Press.
- Prudnikov, A. P., Brychkov, Y. A., and Marichev, O. I. (1990). *Integrals and Series, Vol. 3, More Special Functions*, New York: Gordon and Breach.
- Roberts, S. W. (1966). A Comparison of Some Control Chart Procedures, *Technometrics* 8: 411–430.
- Sharma, J. L. (1938). On Whittaker’s Confluent Hypergeometric Function, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 25: 491–504.
- Shiryaev, A. N. (1961). The Problem of the Most Rapid Detection of a Disturbance in a Stationary Process, *Soviet Mathematics—Doklady* 2: 795–799. Translation from *Doklady Akademii Nauk SSSR* 138: 1039–1042, 1961.
- Shiryaev, A. N. (1963). On Optimum Methods in Quickest Detection Problems, *Theory of Probability and Its Applications* 8: 22–46.
- Shiryaev, A. N. (1978). *Optimal Stopping Rules*, New York: Springer-Verlag.
- Shiryaev, A. N. (1996). Minimax Optimality of the Method of Cumulative Sums (CUSUM) in the Case of Continuous Time, *Russian Mathematical Surveys* 51: 750–751.
- Shiryaev, A. N. (2002). *Quickest Detection Problems in the Technical Analysis of the Financial Data*, in H. Geman, D. Madan, S. R. Pliska, and T. Vorst, eds., *Mathematical Finance—Bachelier Congress 2000*, pp. 487–521, Heidelberg: Springer.

- Shiryaev, A. N. (2006). *From “Disorder” to Nonlinear Filtering and Martingale Theory*, in A. A. Bolibruch, Y. S. Osipov, and Y. G. Sinai, eds., *Mathematical Events of the Twentieth Century*, pp. 371–397, Heidelberg: Springer.
- Shiryaev, A. N. (2011). *Probabilistic–Statistical Methods in Decision Theory*, Yandex School of Data Analysis Lecture Notes, Moscow: MCCME (in Russian).
- Slater, L. J. (1960). *Confluent Hypergeometric Functions*, Cambridge: Cambridge University Press.
- Tartakovsky, A., Nikiforov, I., and Basseville, M. (2014). *Sequential Analysis: Hypothesis Testing and Changepoint Detection*, Boca Raton: CRC Press.
- Tartakovsky, A. G. and Moustakides, G. V. (2010). State-of-the-Art in Bayesian Changepoint Detection, *Sequential Analysis* 29: 125–145.
- Tartakovsky, A. G., Pollak, M., and Polunchenko, A. S. (2012). Third-Order Asymptotic Optimality of the Generalized Shiryaev–Roberts Changepoint Detection Procedures, *Theory of Probability and Its Applications* 56: 457–484.
- Tartakovsky, A. G. and Polunchenko, A. S. (2010). Minimax Optimality of the Shiryaev–Roberts Procedure, in *Proceedings of 5th International Workshop in Applied Probability*, Universidad Carlos III of Madrid, July 5–8, Colmenarejo, Spain.
- Titchmarsh, E. C. (1962). *Eigenfunction Expansions Associated with Second-Order Differential Equations*, Oxford: Clarendon.
- Veeravalli, V. V. and Banerjee, T. (2013). *Quickest Change Detection*, in R. Chellappa and S. Theodoridis, eds., *Academic Press Library in Signal Processing: Array and Statistical Signal Processing*, vol. 3, pp. 209–256. Oxford: Academic Press.
- Whittaker, E. T. (1904). An Expression of Certain Known Functions as Generalized Hypergeometric Functions, *Bulletin of American Mathematical Society* 10: 125–134.
- Whittaker, E. T. and Watson, G. N. (1927). *A Course of Modern Analysis*, 4th edition, Cambridge: Cambridge University Press.
- Wong, E. (1964). The Construction of a Class of Stationary Markoff Processes, in R. Bellman, ed., *Stochastic Processes in Mathematical Physics and Engineering*, pp. 264–276, Providence: American Mathematical Society.